Appendix C: XML Overview

*Introduction to XML**

As a means of providing background knowledge necessary for understanding the terminology discussed in this document the following section offers a synopsis of XML and its layout and structure.

XML stands for e**X**tensible **M**arkup **L**anguage. XML is a recommendation by the World Wide Web Consortium (W3C) for how to represent structured information in a text-based format. XML is a markup language whose roots originate in HTML and SGML. Like HTML, tags and attributes are part of XML syntax. While HTML tags have a specific meaning and generally describe how the text between them should look, XML uses tags only to delimit pieces of data and leaves the interpretation of the data completely to the application. The tags are called elements and they provide metadata or meaning for the contents in between. For example, FirstName is the metadata describing the meaning of Heidi.

> <FirstName>Heidi</FirstName>

<FirstName>is a start tag. FirstName is an element. Heidi is the data, or XML content. </FirstName> is an end tag.

Elements can be simple or complex. Simple elements, like the example above, do not have any sub elements. Complex elements are logical groupings of simple or other complex elements. In the next example, the complex element is in bold.

> **<Name>**
> <FirstName>Heidi</FirstName>
> <LastName>Smith</LastName>
> **</Name>**

Common XML Usage

The textual and self-describing nature of XML makes it platform independent and very flexible in terms of content. Thus, it is an open standard for storing, publishing and exchanging any kind of information between any number of disparate systems. As long as two parties agree on and adhere to a common data structure, they can transfer information to each other without any knowledge of the other party's operating system, database type or programming language.

One if the main features, of XML, is its ability to provide a means for guaranteeing the validity of a document. This is done through the use of XML schemas. An XML Schema specifies the rules surrounding the logical structure of an XML document. It is a language that describes the allowed content of documents. It defines the elements present in the document and the order in

which they appear, as well as any attributes that may be associated with an element. This is a big advantage over other text based methods of communication since errors in these files cannot be discovered until the data contained is being processed.

XML documents can be used as a storage mechanism to hold hierarchical data. These documents can then be queried to extract specific information. This makes XML a powerful tool for content management since data extracted from multiple documents can be re-formatted to produce new documents as required by the application or user. The coupling of XML with Extensible Stylesheet Language (XSL) also allows the same content to be displayed in various formats, from HTML to Braille.

FSA XML Usage

One of the objectives behind designing an XML framework is to provide FSA and our partnering student aid schools greater flexibility in record processing, i.e., opportunities for multiple data cross-walks. XML offers the flexibility to design records, known as XML documents, particular to an audience or community. It allows increased access to and reuse of information. It supports validation [edits] by checking structural validity and flagging errors. It also enables systems to share information and users to see different views of available data.

The following is an example of how a student record might be represented using XML.

```
<Student SSN="299999999" BirthDate="1979-02-03" NameLast="Student">
      <Identifiers>
            <DriversLicense>
                  <DriversLicenseState>IN</DriversLicenseState>
                  <DriversLicenseNumber>DL222222</DriversLicenseNumber>
            </DriversLicense>
      </Identifiers>
      <BirthDate>1979-02-03</BirthDate>
      <Name>
            <FirstName>Sally</NameFirst>
            <MiddleInitial>A</MiddleInitial>
            <LastName>Smith</LastName>
      </Name>
      <Contacts>
            <Address>
                  <AddressFormatIndicator>false</AddressFormatIndicator>
                  <AddressTypeCode>P</AddressTypeCode>
                  <AddressLine>417 Hapler Road</AddressLine>
                  <AddressCity>Fort Wayne</AddressCity>
                  <AddressStateProvinceCode>IN</AddressStateProvinceCode>
                  <AddressPostalCode>46807</AddressPostalCode>
                  <AddressCounty>Allen</AddressCounty>
                  <AddressCountryCode>USA</AddressCountryCode>
```

```
            </Address>
            <PhoneNumber>2197999999</PhoneNumber>
            <EmailAddress>sally.a.smith@email.net</EmailAddress>
        </Contacts>
        <CitizenshipStatusCode>1</CitizenshipStatusCode>
        <NoteMessage>PID=221784902</NoteMessage>
</Student>
```

*XML overview was gathered from COD Technical Reference